Monatsh Math (2009) 156:201–210
DOI 10.1007/s00605-008-0014-x

No cut-off phenomenon for the "Insect Markov chain"

Daniele D'Angeli · Alfredo Donno

Received: 20 April 2007 / Accepted: 26 March 2008 / Published online: 27 August 2008
© Springer-Verlag 2008

Abstract In this work we show that the probability measure associated with the Insect Markov chain defined on the ultrametric space of the leaves of the q-ary rooted tree of depth $n \geq 2$ converges to the stationary distribution without a cut-off behavior.

Keywords Rooted q-ary tree · Ultrametric space · Gelfand pairs · Spherical functions · Spectral analysis · k-Steps transition probability · Cut-off phenomenon

Mathematics Subject Classification (2000) 60J10 · 43A85 · 43A90 · 05C05 · 05C25

1 Introduction

The study of the rate of convergence of an ergodic Markov chain to the stationary distribution has been considered by P. Diaconis in relation with the following question: "How long does it take for an ergodic Markov chain to converge to the stationary distribution π?". This is motivated by the fact that in many Markov chains the distance between the probability measure $m^{(k)}$ determined by the k-steps transition probability and π is close to 0 only after a fixed number k_0 of steps, and it is large (close to 1) before k_0 steps. So the distance exponentially fast breaks down in a small range. This

Communicated by D. Elworthy.

D. D'Angeli (✉) · A. Donno
Section de Mathématiques, Université de Genève, 2-4, Rue du Lièvre,
Case Postale 64, 1211 Geneva 4, Switzerland
e-mail: Daniele.Dangeli@math.unige.ch

A. Donno
e-mail: Alfredo.Donno@math.unige.ch

phenomenon has been called "cut-off phenomenon", a terminology introduced in [1]. See the survey [5] for many applications, or [6] for a new approach to this theory.

In this paper, we show that the Markov chain introduced by A. Figà-Talamanca in [7] does not present this property. The associated Markov chain will be called "insect" (following [4]), and it is defined on the n-th level L_n of the q-ary rooted tree of depth n, denoted by $T_{q,n}$. Each step consists to reach again the n-th level of the q-ary tree after moving "inside" the graph associated with $T_{q,n}$ according with an isotropic random walk. The space L_n is endowed with an ultrametric distance and the probability of reaching a vertex of L_n only depends on the distance from the starting vertex. We study the rate of convergence to the stationary distribution by a spectral analysis. The result is obtained by a direct computation, using some properties of Gelfand pairs theory (for general notions about this topic see, for example [2]). In particular, the k-steps transition probability is expressed in terms of the eigenvalues of the stochastic matrix associated with the Markov chain; they are computed using the spherical functions associated with the action of the full automorphisms group $Aut(T_{q,n})$ on L_n as in [3].

2 Preliminaries

2.1 Basic properties

We start this section by presenting some preliminary results on Markov chains. Our source is [3]. Let X be a finite set. Suppose that P is a reversible transition probability on X, i.e., there exists a strict probability measure π on X such that

$$\pi(x)p(x,y) = \pi(y)p(y,x),$$

for all $x, y \in X$. One says that P and π are in detailed balance.

We can associate with P a Markov operator acting on $L(X) = \{f : X \longrightarrow \mathbb{C}\}$ as $Pf(x) = \sum_{y\in X} p(x,y)f(y)$.

Moreover, the hypothesis of reversibility of P guarantees that P can be diagonalized over $\mathbb{R}$. Let λ_z, for $z \in X$, be the eigenvalues of P. Then we have the following formula for the k-steps transition probability from x to y:

$$p^{(k)}(x,y) = \pi(y)\left(1 + \sum_{z\neq z_0} u(x,z)\lambda_z^k u(y,z)\right), \tag{1}$$

where $u(x,z)_{x,z\in X}$ is a unitary matrix whose columns are eigenvectors for P.

The following definitions are classical.

Definition 2.1 Let $P = (p(x,y))_{x,y\in X}$ be a stochastic matrix. Then a stationary distribution for P is a probability measure π on X such that

$$\pi(y) = \sum_{x\in X} \pi(x)p(x,y),$$

for all $y \in X$.

Definition 2.2 Let $P = (p(x, y))_{x,y \in X}$ be a stochastic matrix. Then P is ergodic if there exists $n_0 \in \mathbb{N}$ such that

$$p^{(n_0)}(x, y) > 0,$$

for all $x, y \in X$.

The following theorem establishes a relation between ergodicity and the spectrum of the operator P.

Theorem 2.3 *Let P be a reversible stochastic matrix on X. Then P is ergodic if and only if the eigenvalue $\lambda_0 = 1$ has multiplicity one and -1 is not an eigenvalue for P.*

Moreover, the following theorem gives a relation between stationary distributions and ergodicity. For a proof see, for example, Chap. 1 in [3].

Theorem 2.4 (Markov Ergodic Theorem) *Let P be a reversible stochastic matrix on X. Then P is ergodic if and only if*

$$\lim_{n\to\infty} p^{(k)}(x, y) = \pi(y) \quad \textit{for all } x, y \in X,$$

where π is the strict probability measure which is in detailed balance with P. Moreover, π is the unique stationary distribution for P.

The next definition will be useful later, because it introduces the notion of distance of two distributions on X.

Definition 2.5 Let μ and ν two probability distributions on X. Then their total variation distance is defined as

$$\|\mu - \nu\|_{TV} = \max_{A\subseteq X} \left| \sum_{x\in A} \mu(x) - \nu(x) \right| \equiv \max_{A\subseteq X} |\mu(A) - \nu(A)|.$$

It is easy to prove that $\|\mu - \nu\|_{TV} = \frac{1}{2}\|\mu - \nu\|_{L^1}$, where

$$\|\mu - \nu\|_{L^1} = \sum_{x\in X} |\mu(x) - \nu(x)|.$$

2.2 The q-ary tree and the associated (finite) Gelfand pairs

Denote by $T_{q,n}$ the q-ary rooted tree of depth n and by L_n its n-th level . Each vertex in L_n can be written as a word $x = x_1 x_2, \ldots, x_n$ in the alphabet $\{0, 1, \ldots, q - 1\}$. The set L_n can be endowed with an ultrametric distance d, defined in the following way: if $x = x_1, \ldots, x_n$ and $y = y_1, \ldots, y_n$, then

$$d(x, y) = n - \max\{i : x_k = y_k, \ \forall\, k \le i\}.$$

We observe that $d = d'/2$, where d' denotes the usual geodesic distance on $T_{q,n}$.

We denote by $Aut(T_{q,n})$ the group of all automorphisms of $T_{q,n}$: if $g \in Aut(T_{q,n})$ the action of g on x is given by

$$g(x) = g_{\emptyset}(x_1)g_{x_1}(x_2), \ldots, g_{x_1x_2,\ldots,x_{n-1}}(x_n),$$

where $g_w \in S_q$ (the symmetric group on q elements) represents the labelling of g at the vertex w, i.e., the restriction of the action of g on the children of w, for every finite word w in the alphabet $\{0, 1, \ldots, q-1\}$.

Let $x_0 = 0^n$ and denote by $K_{q,n} = \{g \in Aut(T_{q,n}) : g(x_0) = x_0\}$ its stabilizer. Then, as the action of $Aut(T_{q,n})$ on L_n is transitive, the homogeneous space $X = Aut(T_{q,n})/K_{q,n}$ can be identified with L_n. The associated pair $(Aut(T_{q,n}), K_{q,n})$ is a Gelfand pair, that is, the subalgebra of bi-$K_{q,n}$-invariant functions in $L(Aut(T_{q,n}))$ (which is isomorphic to the subalgebra of $K_{q,n}$-invariant functions on $L(L_n)$) is commutative (see, for example [3], Chap. 4 for more on Gelfand pairs and Chap. 7 for $(Aut(T_{q,n}), K_{q,n})$) and so it gives rise to the following decomposition of the space $L(L_n)$ into irreducible submodules

$$L(L_n) = \bigoplus_{j=0}^{n} W_j,$$

where

$$W_j = \{f \in L(L_n) : f = f(x_1, x_2, \ldots, x_j) \text{ and } \sum_{x=0}^{q-1} f(x_1x_2, \ldots, x_{j-1}x) \equiv 0\}.$$

In each space W_j there exists a unique function ϕ_j, called a *spherical function*, which is $K_{q,n}$-invariant (that is, $\phi_j(kx) = \phi_j(x)$ for all $k \in K_{q,n}$ and $x \in L_n$) and $\phi_j(x_0) = 1$.

It is well known (see [7]) that, for every $j = 0, 1, \ldots, n$, the spherical function $\phi_j \in W_j$ has the following expression:

$$\phi_j(x) = \begin{cases} 1 & \text{if } d(x, x_0) < n - j + 1 \\ \frac{1}{1-q} & \text{if } d(x, x_0) = n - j + 1 \\ 0 & \text{if } d(x, x_0) > n - j + 1 \end{cases}.$$

Moreover $d_j := \dim W_j = q^{j-1}(q-1)$, for $j = 1, \ldots, n$ and $d_0 := \dim W_0 = 1$.

Finally, we recall that for a Gelfand pair (G, K), with $X = G/K$, the formula (1) becomes (see Chap. 4 in [3])

$$p^{(k)}(x_0, x) = \frac{1}{|X|} \sum_{i=0}^{n} d_i \lambda_i^k \phi_i(x), \tag{2}$$

where $\lambda_i = [\phi_i * p(x_0, \cdot)](1_G)$ is the i-th coefficient of the spherical Fourier transform of $p(x_0, .)$. In other words, $\lambda_0, \lambda_1, \ldots, \lambda_n$ are the eigenvalues of the operator P (of convolution by $p(x_0, .)$).

2.3 Insect Markov chain

In [7] the following Markov chain on the space L_n is defined. Suppose that at time zero we start from the vertex $x_0 = 0^n \in L_n$. Let ξ_i denote the vertex 0^{n-i} and α_i the probability to reach ξ_{i+1} from ξ_i. It is clear that $\alpha_0 = 1$, $\alpha_1 = \frac{1}{q+1}$ and $\alpha_n = 0$. This leads to the following recursive expression

$$\alpha_j = \frac{1}{q+1} + \alpha_{j-1}\alpha_j \frac{1}{q+1}.$$

Solving the equation we get

$$\alpha_j = \frac{q^j - 1}{q^{j+1} - 1}, \quad 1 \leq j \leq n-1.$$

Hence we can define $P = (p(x, y))_{x,y \in L_n}$, as the stochastic matrix whose entry $p(x, y)$ is the probability that y is the first vertex in L_n reached from x in the Markov chain defined above. It is clear that if $d(x, y) = d(x, z)$ (i.e., y and z are in the same ultrametric sphere of center x) we have $p(x, y) = p(x, z)$. Fixed the vertex $x_0 = 0^n$, we can compute, recalling the significance of the α_j's

$$\begin{aligned} p(x_0, x_0) = {} & q^{-1}(1-\alpha_1) + q^{-2}\alpha_1(1-\alpha_2) + \cdots \\ & + q^{-n+1}\alpha_1\alpha_2, \ldots, \alpha_{n-2}(1-\alpha_{n-1}) + q^{-n}\alpha_1\alpha_2, \ldots, \alpha_{n-1}. \end{aligned}$$

It is clear that, if $d(x_0, x) = 1$, then $p(x_0, x) = p(x_0, x_0)$.

More generally, if $d(x_0, x) = j > 1$, we have

$$\begin{aligned} p(x_0, x) = {} & q^{-j}\alpha_1\alpha_2\cdots\alpha_{j-1}(1-\alpha_j) + \cdots \\ & + q^{-n+1}\alpha_1\alpha_2\cdots\alpha_{n-2}(1-\alpha_{n-1}) + q^{-n}\alpha_1\alpha_2\cdots\alpha_{n-1}. \end{aligned}$$

In order to compute the eigenvalues λ_j, $j = 0, 1, \ldots, n$ of the associated operator P one can observe that by the isomorphism between the algebra of $Aut(T_{q,n})$-invariant operators on $L(L_n)$ and the algebra of $K_{q,n}$-invariant functions in $L(L_n)$, it is enough to consider the spherical Fourier transform of the convolver representing P (see [2]), namely

$$\lambda_j = \sum_{x \in L_n} p(x_0, x)\phi_j(x), \quad j = 0, 1, \ldots, n.$$

Using the expressions given for P and the ϕ_j's we get the following eigenvalues.

For $j = 0$, we get

$$\lambda_0 = \sum_{x \in L_n} p(x_0, x) = 1.$$

For $j = n$, we have

$$\lambda_n = p(x_0, x_0) \times 1 + p(x_0, x) \left(-\frac{1}{q-1}\right) \times (q-1) = 0.$$

For $1 \leq j < n$, we get

$$\begin{aligned}
\lambda_j &= qp(x_0, x_1) + (q^2 - q)p(x_0, x_2) + \cdots + (q^{n-j} - q^{n-j-1})p(x_0, x_{n-j}) \\
&\quad + (1-q)^{-1}(q^{n-j+1} - q^{n-j})p(x_0, x_{n-j+1}) \\
&= q(p(x_0, x_1) - p(x_0, x_2)) + q^2(p(x_0, x_2) - p(x_0, x_3)) + \cdots \\
&\quad + q^{n-j-1}(p(x_0, x_{n-j-1}) - p(x_0, x_{n-j})) + q^{n-j}p(x_0, x_{n-j}) \\
&\quad + (1-q)^{-1}(q^{n-j+1} - q^{n-j})p(x_0, x_{n-j+1}) \\
&= \sum_{h=1}^{n-j} q^h (p(x_0, x_h) - p(x_0, x_{h+1})) \\
&= (1 - \alpha_1) + \alpha_1(1 - \alpha_2) + \cdots + \alpha_1\alpha_2, \ldots, \alpha_{n-j-1}(1 - \alpha_{n-j}) \\
&= 1 - \alpha_1\alpha_2, \ldots, \alpha_{n-j} \\
&= 1 - \frac{q-1}{q^{n-j+1} - 1}.
\end{aligned}$$

Observe that P is in detailed balance with the uniform distribution π on L_n given by $\pi(x) = \frac{1}{q^n}$ for all $x \in L_n$. Therefore, after our computations and by virtue of Theorem 2.3, the Insect Markov chain is ergodic.

3 Cut-off phenomenon

3.1 General properties

Let $m^{(k)}(x) = p^{(k)}(x_0, x)$ be the distribution probability after k steps. The total variation distance allows to estimate how $m^{(k)}$ converges to the stationary distribution π.

There are interesting cases in which the total variation distance remains close to 1 for a long time and then tends to 0 in a very fast way (see, for some examples, [5,6]). This suggests the following definition (see [3]).

Consider a sequence (X_n, m_n, p_n), where, for every integer n, X_n is a finite set and m_n, p_n are a probability measure and an ergodic transition probability on X_n, respectively. Denote by π_n the corresponding stationary measure and $m_n^{(k)}$ the distribution of (X_n, m_n, p_n) after k steps.

Now let $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ be two sequences of positive real numbers such that

$$\lim_{n\to\infty} \frac{b_n}{a_n} = 0.$$

Definition 3.1 The sequence of Markov chains (X_n, m_n, p_n) has a (a_n, b_n)-cut-off if there exist two functions $f_1, f_2 : [0, +\infty) \longrightarrow \mathbb{R}$ with

- $\lim_{c\to+\infty} f_1(c) = 0$
- $\lim_{c\to+\infty} f_2(c) = 1$

such that, for any fixed $c > 0$, one has

$$\|m_n^{(a_n+cb_n)} - \pi_n\|_{TV} \leq f_1(c) \quad \text{and} \quad \|m_n^{(a_n-cb_n)} - \pi_n\|_{TV} \geq f_2(c)$$

for sufficiently large n.

The following proposition gives a necessary condition for the cut-off phenomenon.

Proposition 3.2 *If the sequence* (X_n, m_n, p_n) *has an* (a_n, b_n)*-cut-off, then for any* $0 < \epsilon_1 < \epsilon_2 < 1$ *there exist* $k_2(n) \leq k_1(n)$ *such that*

(1) $k_2(n) \leq a_n \leq k_1(n)$;
(2) *for n large,* $k \geq k_1(n) \Rightarrow \|m_n^{(k)} - \pi_n\|_{TV} \leq \epsilon_1$;
(3) *for n large,* $k \leq k_2(n) \Rightarrow \|m_n^{(k)} - \pi_n\|_{TV} \geq \epsilon_2$;
(4) $\lim_{n\to\infty} \frac{k_1(n)-k_2(n)}{a_n} = 0$.

Proof By definition there exist c_1 and c_2 such that $f_2(c) \geq \epsilon_2$ for $c \geq c_2$ and $f_1(c) \leq \epsilon_1$ for $c \geq c_1$. So it suffices to take $k_1(n) = a_n + c_1 b_n$ and $k_2(n) = a_n - c_2 b_n$ to get the assertion. □

3.2 The case of Insect Markov chain

The cut-off phenomenon occurs in several examples of Markov chains. In general it can be detected thanks to a careful spectral analysis, as we will do in the proof of the following theorem. In what follows suppose $n \geq 2$.

Theorem 3.3 *The probability measure associated with the Insect Markov chain converges to the stationary distribution without a cut-off behavior.*

Proof We want to give an expression for $m^{(k)}(x) = p^{(k)}(x_0, x)$. From (2) we get

- If $x = x_0$, then

$$m^{(k)}(x_0) = \frac{1}{q^n} \left\{ 1 + \sum_{j=1}^{n} q^{j-1}(q-1) \left[1 - \frac{q-1}{q^{n-j+1}-1} \right]^k \right\}.$$

- If $d(x_0, x) = h$, with $1 \leq h \leq n-1$, then

$$\begin{aligned} m^{(k)}(x) &= \frac{1}{q^n} \left\{ 1 + \sum_{j=1}^{n-h+1} q^{j-1}(q-1) \left[1 - \frac{q-1}{q^{n-j+1}-1} \right]^k \phi_j(x) \right\} \\ &= \frac{1}{q^n} \left\{ 1 + \sum_{j=1}^{n-h} q^{j-1}(q-1) \left[1 - \frac{q-1}{q^{n-j+1}-1} \right]^k - q^{n-h} \left[1 - \frac{q-1}{q^h-1} \right]^k \right\} \end{aligned}$$

- If $d(x_0, x) = n$, then

$$m^{(k)}(x) = \frac{1}{q^n}\left\{1 - \left[1 - \frac{q-1}{q^n - 1}\right]^k\right\}.$$

Let π be the uniform distribution on L_n. Then we have

$$\|m^{(k)} - \pi\|_{L^1} = \frac{1}{q^n}\left\{\sum_{j=1}^{n} q^{j-1}(q-1)\lambda_j^k \right.$$
$$+ \sum_{h=1}^{n-1}(q^h - q^{h-1})\left|\sum_{j=1}^{n-h} q^{j-1}(q-1)\lambda_j^k - q^{n-h}\lambda_{n-h+1}^k\right|$$
$$\left. + q^{n-1}(q-1)\lambda_1^k\right\}.$$

Now observe that

$$\frac{1}{q^n}\sum_{h=1}^{n-1}(q^h - q^{h-1})\sum_{j=1}^{n-h} q^{j-1}(q-1)\lambda_j^k + \frac{1}{q^n}\sum_{j=1}^{n} q^{j-1}(q-1)\lambda_j^k$$
$$= \frac{1}{q^n}\sum_{j=1}^{n-1}\left[1 + (q-1) + (q^2 - q) + \cdots + (q^{n-j} - q^{n-j-1})\right]\cdot q^{j-1}(q-1)\lambda_j^k$$
$$= \frac{1}{q^n}\sum_{j=1}^{n-1} q^{n-1}(q-1)\lambda_j^k = \frac{q-1}{q}\sum_{j=1}^{n-1}\lambda_j^k$$

and

$$\frac{1}{q^n}\sum_{h=1}^{n-1}(q^h - q^{h-1})q^{n-h}\lambda_{n-h+1}^k + \frac{1}{q^n}(q^n - q^{n-1})\lambda_1^k = \frac{q-1}{q}\sum_{j=1}^{n-1}\lambda_j^k.$$

Using the trivial fact that $\sum_j |a_j - b_j| \leq \sum_j (|a_j| + |b_j|)$, we conclude

$$\|m^{(k)} - \pi\|_{L^1} \leq \frac{2(q-1)}{q}\sum_{j=1}^{n-1}\lambda_j^k.$$

On the other hand

$$\|m^{(k)} - \pi\|_{L^1} \geq \sum_{x: d(x_0, x) = n} |m^{(k)}(x) - \pi(x)|$$
$$= \frac{1}{q^n}(q^n - q^{n-1})\lambda_1^k = \frac{q-1}{q}\lambda_1^k.$$

So we get the following estimate:

$$\frac{q-1}{q}\lambda_1^k \leq \|m^{(k)} - \pi\|_{L^1} \leq \frac{2(q-1)}{q}\sum_{j=1}^{n-1}\lambda_j^k,$$

or, equivalently,

$$\frac{q-1}{2q}\lambda_1^k \leq \|m^{(k)} - \pi\|_{TV} \leq \frac{(q-1)}{q}\sum_{j=1}^{n-1}\lambda_j^k.$$

In what follows the following inequalities will be used:

(1) $(1-x)^k \leq \exp(-kx)$ if $x \leq 1$.
(2) $\frac{q^n-1}{q^{n-j+1}-1} \geq q^{j-1}$, for $j \geq 1$.
(3) $q^{j-1} \geq j$, for $q \geq 2$ and $j \geq 1$.

Choose $k_2(n) = \frac{q^n-1}{q-1}$, then (recall the computations at page 5)

$$\begin{aligned}
\frac{q-1}{q}\sum_{j=1}^{n-1}\lambda_j^k &\leq \frac{q-1}{q}\sum_{j=1}^{n-1}\exp\left(-\frac{q-1}{q^{n-j+1}-1}k\right) \leq (\text{ if } k \geq k_2(n))\\
&\leq \frac{q-1}{q}\sum_{j=1}^{n-1}\exp\left(-\frac{q-1}{q^{n-j+1}-1}k_2(n)\right)\\
&\leq \frac{q-1}{q}\sum_{j=1}^{n-1}\exp(-q^{j-1}) \leq \frac{(q-1)}{q}\sum_{j=1}^{n-1}(e^{-j})\\
&\leq \frac{(q-1)}{q}\sum_{j=1}^{\infty}(e^{-1})^j = \frac{q-1}{q}\cdot\frac{1}{e-1} := \epsilon_2.
\end{aligned}$$

On the other hand, if $k_1(n) = 2\frac{q^n-1}{q-1}$ $(= 2k_2(n))$, we get

$$\begin{aligned}
\frac{q-1}{2q}\lambda_1^k &= \frac{q-1}{2q}\left[1-\frac{q-1}{q^n-1}\right]^k \geq (\text{ if } k \leq k_1(n))\\
&\geq \frac{q-1}{2q}\left[1-\frac{q-1}{q^n-1}\right]^{2\frac{q^n-1}{q-1}} \geq \frac{q-1}{2q}e^{-3} := \epsilon_1.
\end{aligned}$$

Now $k_1(n) > k_2(n)$, $\epsilon_1 < \epsilon_2$ and

- for $k \geq k_2(n)$ we have $\|m^{(k)} - \pi\|_{TV} \leq \epsilon_2$,
- for $k \leq k_1(n)$ we have $\|m^{(k)} - \pi\|_{TV} \geq \epsilon_1$.

This implies that cut-off phenomenon does not occur in this case by Proposition 3.2. In fact, the sequences $k_1(n)$ and $k_2(n)$ cannot satisfy condition (4) of Proposition 3.2. This gives the assertion. □

Remark 3.4 Using the same strategy of Theorem 3.3 one can easily check that the cut-off phenomenon does not occur also if we fix n and let $q \to +\infty$.

Remark 3.5 If $n = 1$ we get the simple random walk on the complete graph K_q on q vertices, in which each vertex has a loop. It is straightforward that the first step is performed by equiprobably choosing anyone of the q vertices and so the probability measure $m^{(1)}$ already equals the uniform distribution π on the set of the vertices.

References

1. Aldous, D., Diaconis, P.: Shuffling cards and stopping times. Am. Math. Monthly **93**, 333–348 (1986)
2. Ceccherini-Silberstein, T., Scarabotti, F., Tolli, F.: Finite Gelfand pairs and their applications to probability and statistics. J. Math. Sci. NY. **141**(2), 1182–1229 (2007)
3. Ceccherini-Silberstein, T., Scarabotti, F., Tolli, F.: Harmonic analysis on finite groups: representation theory, Gelfand Pairs and Markov chains. In: Cambridge Studies in Advanced Mathematics, vol. 108. Cambridge University Press, Cambridge (2008)
4. D'Angeli, D., Donno, A.: Crested products of Markov chains. Ann. Appl. Probab. (accepted) (2008)
5. Diaconis, P.: The cut-off phenomenon in finite Markov chains. Proc. Nat. Acad. Sci. USA. **93**, 1659–1664 (1996)
6. Diaconis, P., Saloff-Coste, L.: Separation cut-offs for birth and death chains. Ann. Appl. Probab. **16**, 2098–2122 (2006)
7. Figà-Talamanca, A.: An application of Gelfand pairs to a problem of diffusion in compact ultrametric spaces. In: Topics in Probability and Lie Groups: Boundary Theory, CRM Proc. Lecture Notes, vol. 28, Am. Math. Soc. Providence, RI, pp. 51–67 (2001)